

Interests and Incentives in Government Peer Review

**Prefatory Remarks for
the Society for Risk Analysis Forum on Peer Review
May 29, 2002
Arlington, VA**

**Richard B. Belzer
President
Regulatory Checkbook
Washington, DC
Belzer@RegulatoryCheckbook.org
202.898.2050**

Some (but not all) of the complaints that have been leveled against government peer review reflect legitimate problems with its implementation by federal agencies. Others, however, actually relate to designed-in features and thus are not appropriately construed as implementation problems. Still others bemoan phenomena that are simply not reducible unless and until a race of scientists can be created whose members lack all human qualities, and until policy decisions can and ought to be made by automaton.

In this presentation, I will analyze government peer review not to complain or criticize it, but rather to identify those salient features that are the source of our collective misery. My analytical approach is positive (*i.e.*, descriptive) economics rather than moral philosophy (*i.e.*, prescriptive). Government peer review performs much the way one should expect it to perform given the interests of the people, institutions and issues involved. Improving its performance requires that we have fundamental clarity about three things:

1. what we reasonably expect peer review to accomplish;
2. whether we are all agreed that these objectives are so reasonable and appropriate that we are all willing to commit ourselves *ex ante* to peer review as a process-based management approach for complex science-policy disputes; and
3. whether we can design a peer review system that imparts behavioral incentives that are consistent with our objectives and our commitment to process.

If any of these items is lacking, we cannot expect significant improvement and we can expect further deterioration.

IS PEER REVIEW THE RIGHT TOOL FOR THE PROBLEM?

The answer clearly depends on what we are trying to accomplish, and much disagreement about peer review arises from lacking a shared consensus on goals and objectives. Unlike many others who seek to achieve a meeting of the minds on this

matter, I believe that true agreement is unlikely and may be impossible. The stakes are too high; and the interests of parties interested in peer review are too divergent, discordant and conflicted to reach any genuine *ex ante* consensus and *ex post* consensus is unimaginable in a pluralistic democracy. Even if all the relevant players were here at this meeting and a common written statement could be agreed upon today, consensus would begin to unravel before the first signatory left the room as each of us plotted to reinterpret other views out of the agreement and make our views transcendent.

Peer review is a tool borrowed from academic and scholarly settings. For many of us, our first real introduction to “peer review” occurred during the writing and defending of our doctoral dissertations. We had committees of reviewers led by autocratic chairmen – benign autocrats, of course, if they are still alive and we want to remain on speaking terms. We chose the members of our committees and tried very hard to please and persuade them, for they held all the cards. We were not finished until our “peer reviewers” said we were.

Subsequently most of us began applying for grants and shopping our scholarly wares in front of journal editors.¹ We promised whatever grantors said they wanted and we accommodated any tomfool thing an editor demanded to get our manuscripts accepted. Grantors and editors were the new autocrats, but peer reviewers were anonymous and accountable only to the grantors and editors for whom they allegedly toiled. For some of us, the path to publication was eased by prior success and perhaps quiet rewards for service as peer reviewers and on editorial boards.

The fundamental question raised by journal editors is, “Does this body of work deserve to be published given the limited number of pages available and the quality of competing manuscripts?” It is not that different from our dissertation committee’s fundamental question: “Does this thesis candidate’s work meet appropriate quality standards for our institution and program?”

At no point during that debate does scholarly peer review ask, “Is this manuscript essentially correct in all its data, methods, inferences and conclusions?” That’s a matter for Science to figure out, and Science takes its own sweet time. However, this is precisely the fundamental question asked by governmental peer review. Government peer review devotes little attention to whether a document *ought* to be published, because the reviewee-agency doesn’t need peer reviewers’ permission to publish and surely isn’t constrained by page limitations.²

¹ I myself did not follow this track, accepting a position as staff economist in the Office of Information and Regulatory Affairs at the Office of Management and Budget. I did not apply for grants because federal regulatory agencies kept me supplied with work peer reviewing draft regulations and regulatory impact analyses. I did not seek publication of my work because the working motto at OMB is was “publish, then perish.”

² Agencies may soon be constrained somewhat by information quality limitations, as new law and government-wide guidance has been enacted to deter the use of low-quality science.

Additional problems arise to the extent that hidden objectives truly dominate. A reviewee-agency might care deeply about the state of scientific knowledge or the fundamental truthfulness of its document under review. It also might not. It might care more about whether science can justify what it wants to say and do. It also might be keenly interested in whether the responsibility for controversial or unpleasant policy decisions can be implicitly delegated to others, such as Scientists.

DIFFERENCES IN THE RELATIONSHIP BETWEEN REVIEWEE AND REVIEWER

In academic and scholarly settings, the reviewee is best understood as a supplicant. The thesis chairman is superior to other committee members who are superior to the candidate. Power rests predominantly with the chairman, who along with other committee members performs functions that are part of their normal academic responsibilities. There are no deadlines. The process is private but the reviewers are known exactly. Nothing can force these people to be cooperative, complacent, speedy, encouraging, supportive or polite. Depending on the school, chocolate chip cookies or Alice B. Toklas brownies sometimes help.

The authors of manuscripts submitted to scholarly journals face a similar (but not identical) relationship. They are still supplicants. Editors obtain peer reviewers through a variety of means, but authors do not select them and they remain anonymous.³ They can (and sometimes literally do) take forever to complete a review. There are very strict deadlines that nobody pays real attention to. All power rests with editors.

Government peer review is so different it ought not be called by the same name. The reviewee-agency is not a supplicant but a client. In many cases, the reviewee-agency directly selects the peer reviewers. In other cases, a firewall of uncertain temperature resistance separates the reviewee from those who formally select the reviewers. Typically, the reviewee-agency retains the authority to veto selections made by others, which is almost equivalent to enjoying the power to select. Reviewee-agencies write the charge, fund and staff the process, control the flow of information, and manage the discussion agenda if the panel chairman lacks a firm hand.

Government peer reviewers have deadlines. They are publicly known and are directed to do their work in a fishbowl. Frequently the staff of the reviewee-agency gets an early and private peek at the peer reviewers' work. This would be considered an unethical practice in scholarly peer review.

³ Anonymity tends to become asymmetrical as scholars become more specialized and experienced. Authorship cannot be disguised in many cases because the identity of those working on specific issues or projects is widely known. Peer reviewers can retain anonymity by writing their reviews strategically.

THE DOMINANT STRUCTURAL PARADIGM FOR GOVERNMENT PEER REVIEW

In theory, government peer review is supposed to have a stack of certain attributes. It is supposed to be external to the reviewee-agency and independent from it, comprised of members who are expert in those scientific disciplines relevant to the task at hand, free of conflicts of interest that could undermine their impartiality, focused on science, transparent to public observation and scrutiny, and intellectually vigorous. A succinct version of this structure can be found in OMB's guidance to agencies on Executive Order 12866 implementation issued on September 20, 2001. Federal agencies that follow peer review processes containing these attributes are said to gain deference when their work is ultimately reviewed by my former colleagues at OMB.

It is too early to say whether OMB actually will confer that deference, though my view is that it won't have any other option. The problem is that there is a substantial discrepancy between theory and practice on the matter of whether government peer review structures actually have these attributes.

External and independent. "External" and "independent" are not synonyms. Only employees of the reviewee-agency are excluded from service on an "external" peer review panel. While this surely makes sense, it is not a particularly demanding requirement. "Independence" is far more difficult to achieve. Some peer reviewers are only nominally independent (or, for that matter, nominally external) because their employers depend critically on financial support, grants and contracts from the reviewee-agency. Typically, "independence" is presumed to include only a financial margin when *intellectual* independence may be much more relevant. An external peer reviewer who is fully concordant with a reviewee-agency's policy views, for example, offers precious little true independence.

Independence is always constrained insofar as peer reviewers are explicitly or implicitly selected by the reviewee-agency, which should not be expected to aggressively search for and appoint its sharpest critics. In many instances, that is precisely what a grantor or journal editor might do to ensure that a highly controversial proposal or manuscript receives the most rigorous of reviews. Reviewee-agencies should be expected to avoid such peer reviewers to the maximum extent possible.

Less well understood is the problem associated with one-time versus repeated transactions. A peer reviewer who is professionally flattered by being selected but knows in advance that she will not be asked to serve again is more inclined to act like a juror in a civil trial. This probably enhances objectivity because it removes from the equation any sense that discretion might be rewarded by a return engagement. Conversely, a peer reviewer who earnestly desires to return will be sorely tempted to pull his punches.

This is an ironic twist on the longstanding notion in economics that efficiency and fair dealing are enhanced by repeated transactions. These are memorialized in the adage "the customer is always right." Sellers strive to keep customers happy by promising to resolve

problems even if they are fully customers' own fault. Markets with repeat transactions stimulate socially desirable strategic behavior. Markets dominated by one-time or rare transactions (e.g., buying a rebuilt automobile motor, hosting weddings and funerals) are the most prone to fraud and other shady practices. That is not necessarily the case in government peer review, where strategic behavior on the part of reviewee-agencies and peer reviewers may diminish the quality of peer review. The prospect of a future relationship can motivate a peer reviewer to accommodate the reviewee-agency's interests, and a reviewee-agency to cater to a host of relatively unimportant whims of peer reviewers.

Expertise. The presence of expertise on government peer review panels is rarely a problem, but oftentimes the expertise of the members is poorly matched to the issue at hand. This may occur for any number of reasons. The "most" expert candidates may be unavailable or unwilling to serve. Those selecting the panel may not fully understand what is needed, or the panel itself may discover that it lacks expertise that it didn't know it would need. Peer reviewers (especially economists!) may be unenthusiastic about serving for nominal pay and maximal headaches. It is still the case that you get what you pay for, so the absence of financial compensation has to be made up elsewhere – in prestige, in repeat engagements, or perhaps in private and privileged access in the halls of the reviewee-agency.

In adequate or poorly targeted expertise also may be the product of strategic behavior on the part of the reviewee-agency (or other party with authority to select members). A gap in expertise may be intended if the panel selector -agency wants to skirt a particularly difficult issue. Frequently, there is but one expert in any specific and important issue on a panel. The absence of competing expertise creates the potential for substantially diminished intellectual vigor, as members defer to each other in areas where they perceive another panelist having superior knowledge. (Deferring to other members also may motivate them to defer to you – an excellent side benefit.)

A relatively new phenomenon that threatens to severely undermine expertise (as well as several other desirable attributes) is the selection of peer reviewers because of the stakeholders they represent rather than the expertise they bring. My view is a stark and uncompromising one: stakeholders belong on policy advisory committees but not on peer review panels. If peer review panels are slow slowly transformed into stakeholder groups, then science will be a sure casualty.

Conflicts of interest. Our infatuation with conflict-of-interest seems to be legalistic, misapplied and ultimately self-defeating. As constructed in law, conflict-of-interest has almost nothing to do with improper or unethical conduct (which is clearly proscribed but well nigh impossible to observe) and everything to do with appearances (which can be readily observed at any level of abstraction one might choose). Thus, a peer reviewer may be considered "conflicted" not because of anything she did or said, or because of any hint of partiality in her scientific analysis, but because she is affiliated with an organization that has a financial interest in the outcome.

Moreover, this tenuous definition of conflict is almost always limited to for-profit financial interests. University professors are exempt because their employers are nonprofits and their financial interests, however memorable, are laundered through their universities' sponsored research programs. Noneducational nonprofits (such as mine) are presumptively free of conflicts because nonprofit status confers a façade of social beneficence to which no for-profit enterprise is entitled irrespective of the relative social value of their work. The unalterable fact is that almost all nonprofits (including mine) attract funds because of what we say and do, and our choice of what to say and do influences who is interested in contributing. The only nonprofits that are free of this kind of conflict of interest are the endowed foundations whose financial positions are totally unaffected by the speech and conduct of their leaders. Ironically, such foundations are oftentimes bastions of extremely intense and luridly corrupting interest if their employees are permitted to serve on peer review panels.

Conventional discourse on conflict-of-interest is limited to financial matters, which in the case of scientific peer review may actually be relatively unimportant. The critical kind of interest that matters in peer review is *intellectual* interest, and *conflicts* of intellectual interest are generally desirable. It is only through conflicts of intellectual interest that peer review can sort out competing viewpoints and provide practical utility to the public. A peer review panel that lacks conflicts of intellectual interest is a stultifying enterprise whose primary output is likely to be mere conformity to common dogma.

One final note on conflict of interest: too much attention seems to be devoted to what economists call *inframarginal* concerns. For example, an expert who is employed in a scientific capacity by a giant, multinational company can have but a trivial effect on his share of the company's success or failure through service on a government peer review panel. In contrast, an otherwise identical expert who is a self-employed consultant can find her firm thrives or withers depending on her service as a peer reviewer. In the case of the scientist working for the giant company, all his efforts lie below the margin because they have little or no detectable effect on his livelihood. In the case of the self-employed consultant, however, virtually anything she does influences the fortunes of her company. Similar lessons apply to university professors and other "nonprofit" actors.

THERE IS MUCH MORE THAN SCIENCE AT STAKE IN GOVERNMENT PEER REVIEW

It is a remarkably resilient conceit that government peer review is about Science. It is not. It is about policy. If it weren't about policy, we wouldn't care so much. When agency officials say that they are just following the advice of Scientists or they really want the input of Scientists before making a controversial decision, I consider it no different than the megamillionaire athlete who says his contract dispute with his team is "not about money." That which is the stated reason is least likely to be true, and that which is denied is the most plausible explanation.

Sometimes science illuminates alternative policy choices and makes decisions easier or less controversial. Sometimes. More often, it seems, science is used as a means toward

policy ends. The tool of peer review is abused when reviewee-agencies seek to convert policy disputes into scientific arcana. This tactic misleads the public and diminishes the reputation of the scientific enterprise. It is especially disturbing to see agency officials abdicate their responsibility for making hard decisions by enlisting a peer review panel to provide scientific cover.

Scientists are (mostly) human. When invited to opine on policy matters, scientists are generally inclined to do so with the same regularity and intensity of others. When tempted to craft public policy under the guise of Science, scientists can easily fall prey to the lures of self-importance and opportunity. While scientists have no special expertise in policy matters, their voices gain in relative strength when they become implicit policy or decision makers.

Transparency. Academic and scholarly peer review – the setting from which this tool was adapted for government use – is notoriously opaque. Peer reviewers have been known to invent scientific objections to papers they simply didn't understand. In contrast, government peer review is supposed to be fully transparent. The academics and scholars who are brought in to be government peer reviewers are not at all accustomed to conducting peer review in a fishbowl. Some handle it with equanimity; some grandstand; others withdraw into scholarly shells.

Much of the work of government peer reviewers cannot be done in public, of course. There is too much to do with too little time; the settings are not conducive to rigorous analysis and scholarly reflection; and difficult and delicate questions that can be resolved responsibly in private may never be raised if all communications have to be public. This means government peer review is knowingly and perhaps purposefully schizophrenic. It is only as transparent as law, custom or public demands require. It is, therefore, only superficially transparent.

Vigor. If vigorous effort and examination are essential attributes of government peer review, it is rather amazing how many obstacles lay in the way. First, government peer review systems are designed to maximize collegiality and consensus. These attributes are consistent with satisfying a client reviewee-agency but they are foreign to scholarly peer review whence the peer review model came. Second, peer review tasks are often so complex that there is little or no overlap among reviewers in relevant expertise. Lack of overlap reduces effective expertise to one or two reviewers on any issue. That makes a panel – even a large one – less of a panel and more a collection of individual experts. This is not a formula for vigorous intellectual discourse. Third, government peer review systems tend to provide a dominant role for the reviewee-agency. If the subject of the peer review controls panel selection, writes the charge, sets the schedule for meeting and report deadlines, staffs the panel, manages document flow, and many times actually runs the meetings, the resulting peer review cannot be expected to be vigorous.

REMEDIES

Government peer review is designed and operated to achieve conflicting goals under trying conditions with personnel who may not be properly equipped to do a job they were asked to do that might well be different its actual purpose. It's no wonder we have problems.

Identifying effective remedies depends first on reestablishing clarity of purpose. This means drawing clear distinctions between stated and revealed preferences and objectives, and designing reforms that penalize gaps and reward transparent coincidence. The interests (whether conflicting or coincident) of all interested parties then must be properly understood without euphemism or favor, because understanding these interests is essential for designing a system that has incentives which encourage the cultivation of desirable process attributes. The final step is to identify these specific process reforms and suggest how any why they would bring government peer review closer to its aspirations.

During the Forum, I will identify and elaborate on a number of specific reform proposals. These proposals include structural changes to overcome inherent structural flaws, and functional changes to overcome dysfunctional behavior. I will elaborate at length on the fundamental problem of conflict-of-interest – not the exhibition of actual bias and partiality by government peer reviewers, but the unhealthy fixation on legalistic rules and perceived conflicts that capture only a small portion of real problems, and tend to do so rather badly. Unless this particular problem is addressed, government peer review appears to be well on the way to becoming fatally bollixed in useless, paralyzing procedure.